Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Verma S, Miles D, Gianni L, et al. Trastuzumab emtansine for HER2-positive advanced breast cancer. N Engl J Med 2012;367:1783-91. DOI: 10.1056/NEJMoa1209124

Supplementary Appendix

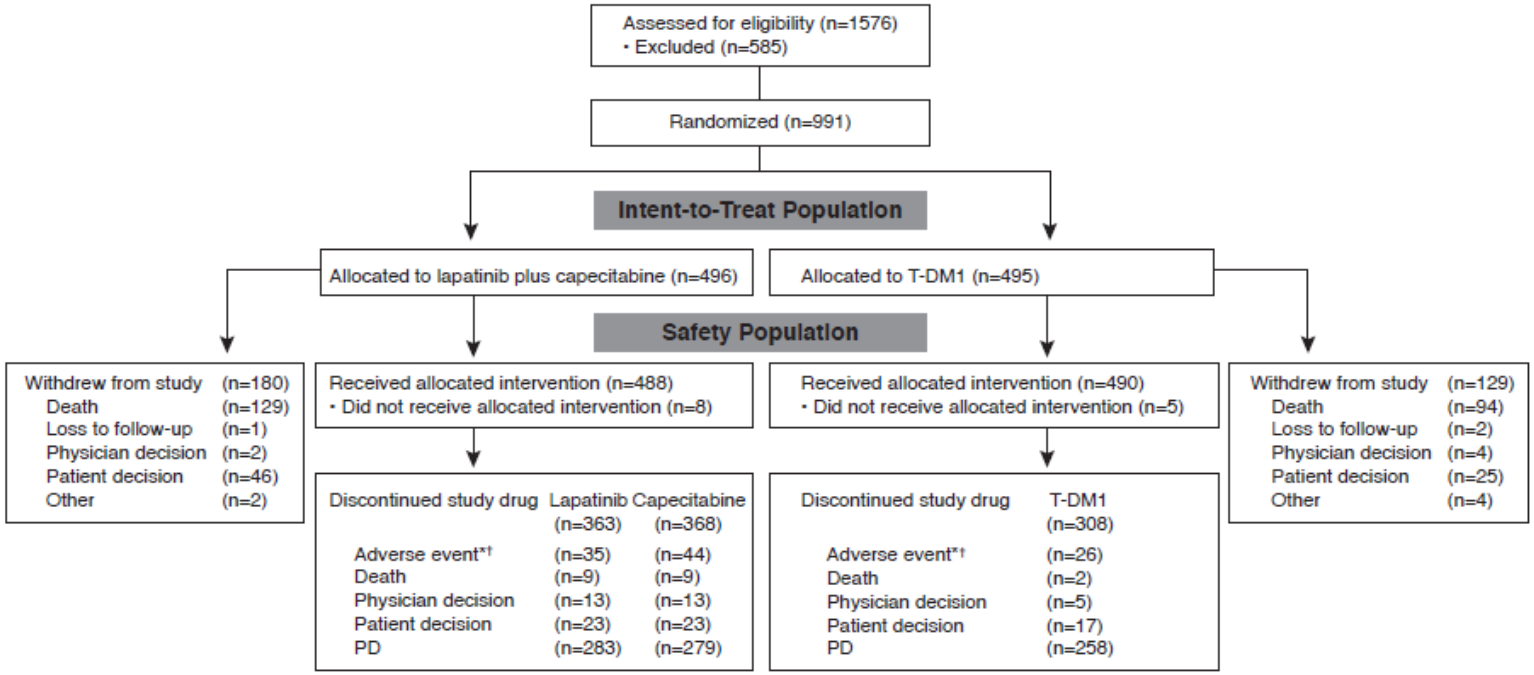
Supplement to: Verma S, Miles D, Gianni L, et al. Trastuzumab Emtansine for HER2-positive Advanced Breast Cancer

CONTENTS:

The following investigators participated in the EMILIA study:

Bosnia and Herzegovina – S Beslija, Z Gojkovic; *Brazil* – G Amorim, RM Araujo, G Borges, R Freitas Jr, R Hegg, JL Pedrini, R Pereira, H Pinczowski, M Portella, VL Silva Teixeira, J Vinholes; *Bulgaria* – A Dudov, V Taskova, C Timcheva, A Tomova; *Canada* – S Ellard, K Gelmon, W Miller, C Prady, A Robidoux, A Robinson, A Scarfe, S Verma, Sh Verma, M Webster, TH Younis; *Denmark* – S Cold; *Finland* – T Alanko, R Huovinen, M Tanner; *France* – D Allouache, M Campone, B Coudert, M Debled, T Delozier, V Dieras, J Grenier, A-C Hardy-Bessard, E Luporsi, F Priou, G Romieu, H Simon; *Germany* – H Eidtmann, K Friedrichs, G Graffunder, V Hagen, G Heinrich, C Hielscher, P Klare, YD Ko, E Stickeler, O Tome, M Untch, M Welslau; *Hong Kong* – J Tsang; *India* – V Raina; *Italy* – D Amadori, G Bianchi, C Bighin, C Boni, F Cognetti, A Contu, D Crivellari, G Curigliano, S de Placido, L del Mastro, A Falcone, L Gianni, A Martoni, F Montemurro, F Roila, A Santoro, M Venturini; *Korea* – SA Im, Y-H Im, SH Lee, D-Y Oh, J Ro, J-H Sohn; *Mexico* –M Gaytan Angel, CA Hernandez, G Tellez-Trevilla; *New Zealand* – R Isaacs; *Philippines* – PF Pua; *Poland* – K Drosik, J Jassem, T Pienkowski, E Staroslawska, R Szoszkiewicz, M Wojtukiewicz, J Zaluski; *Portugal* – N Afonso, MI Pazos; *Russia* – V Semiglazov, O Zharkova; *Singapore* – SC Lee, R Ng; *Slovenia* – S Borstnar; *Sweden* – J Ahlgren, L Klint, J Yachnin; *Switzerland* – J Huober, R Winterhalder; *Spain* – A Anton Torres, J Baselga, E Ciruelos Gil, MA Climent Duran, P Gomez Pardo, A LLombart, S Morales Murillo, A Redondo Sanchez, M Ruiz Borrego; *Taiwan* – H-K Chang, M-C Liu, K-M Rau, H-C Wang, D-C Yeh; *UK* – A Armstrong, PJ Barrett-Lee, M Beresford, J Bishop, A Chakrabarti, P Ellis, T Hickish, D Miles, E Sims, N Turner, M Verrill; *USA* – H Ali, R Agajanian, J Baar, D Berd, K Blackwell, R Blanchard, P Blanchet, R Boccia, B Bower, A Brufsky, T Budd, D Budman, A Bulgaru, D Chan, J Charlson, S Chennuru, S Chitneni, N Christiansen, ML Citron, M Cobleigh, W Conkright, T Cosgriff, S Dakhil, B DiCarlo, J Eneman, L Fehrenbacher, R Frank, E Gartner, L Gressot, J Goldberg, T Guthrie Jr, A Hassan, L Hart, BA Hellerstedt, T Helsten, D Henry, R Hermann, S Hurvitz, N Iannotti, A Kashyap, J Knost, IE Krop, A Kudelka, J Leach, S Limentani, KMS Lo, D Medgyesy, L Meza, S Modi, M Moezi, T Moore, S Murali, J Neidhart, I Oliff, R O'Regan, R Orlowski, S Papish, R Parikh, R Patel, A Perez, J Polikoff, T Pluard, D Riseberg, S Roshon, C Ruud, M Saleh, I Sanchez, F Senecal, P Seo, R Somer, G Somlo, M Shtivelband, S Sundaram, S Swain, H Tezcan, J Wade, S Waintraub, J Waisman, JM Wallmark, T Walters, G Wang, M Wilkinson, D Yardley

Figure S1. Enrollment, Intent-to-Treat and Safety Populations, Treatment Discontinuations, and Withdrawals at the Time of the Progression-Free Survival Analysis. PD denotes progressive disease, T-DM1 trastuzumab emtansine.



*Two patients in the lapatinib + capecitabine arm and three patients in the T-DM1 arm had both an adverse event and progressive disease at the time of treatment discontinuation, with progressive disease attributed as the primary reason for discontinuation.

†The most common adverse events leading to lapatinib or capecitabine discontinuation were diarrhea (n=12) and vomiting (n=11), and diarrhea (n=14), respectively. The most common adverse event leading to T-DM1 discontinuation was thrombocytopenia (n=10).

Figure S2. Progression-free Survival by Independent Review in Patient Subgroups. The vertical dashed line indicates the HR for all patients. Cap denotes capecitabine, CI confidence intervals; ECOG Eastern Cooperative Oncology Group; ER estrogen receptor; HR hazard ratio; LABC locally advanced breast cancer; Lap lapatinib, MBC metastatic breast cancer; PR progesterone receptor, T-DM1 trastuzumab emtansine.

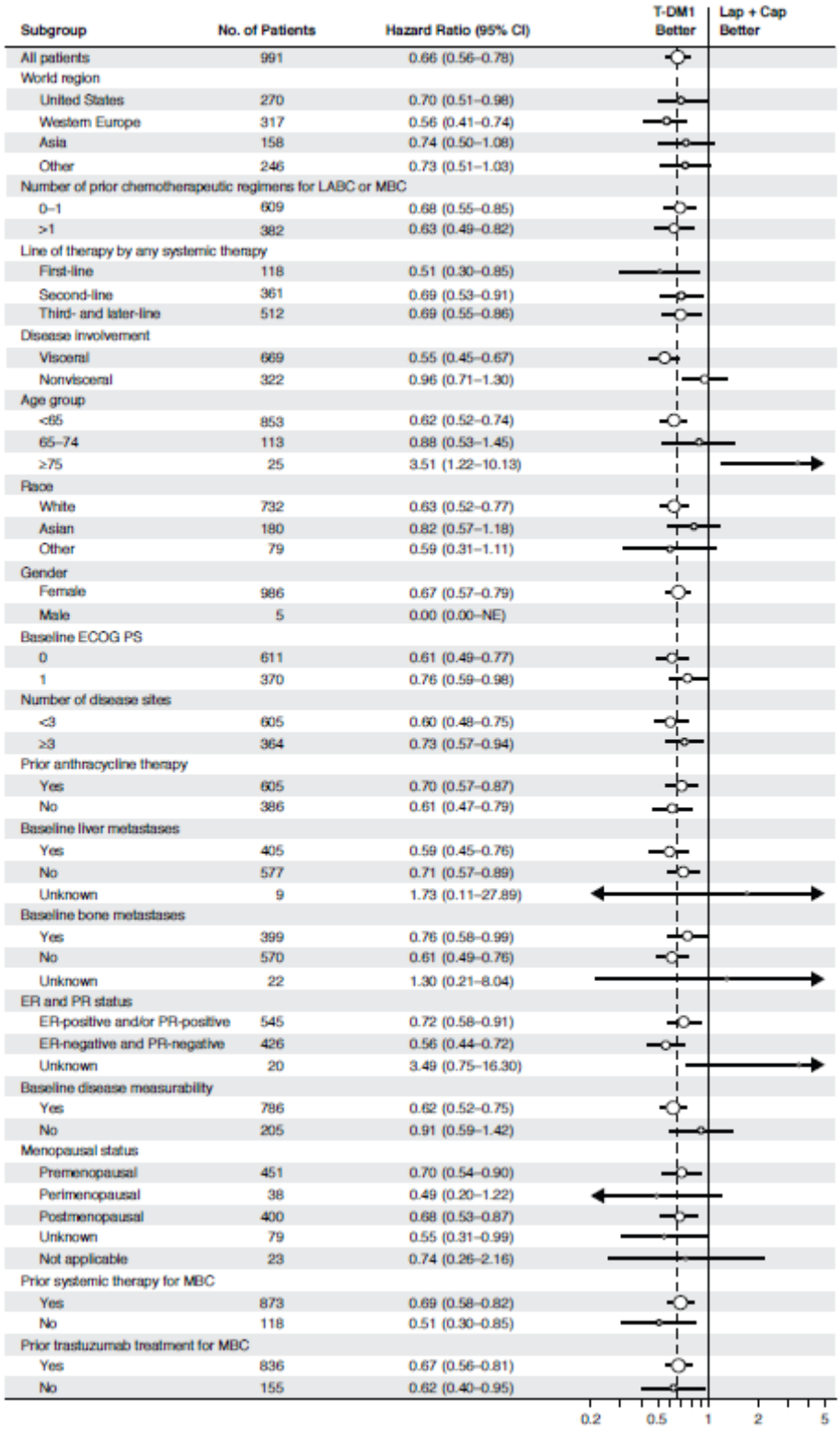

Subgroup
No. of Patients
Hazard Ratio (95% CI)
T-DM1 Better
Lap + Cap Better
All patients 991 0.66 (0.56–0.78)
World region
United States 270 0.70 (0.51–0.98)
Western Europe 317 0.56 (0.41–0.74)
Asia 158 0.74 (0.50–1.08)
Other 246 0.73 (0.51–1.03)
Number of prior chemotherapeutic regimens for LABC or MBC
0–1 609 0.68 (0.55–0.85)
>1 382 0.63 (0.49–0.82)
Line of therapy by any systemic therapy
First-line 118 0.51 (0.30–0.85)
Second-line 361 0.69 (0.53–0.91)
Third- and later-line 512 0.69 (0.55–0.86)
Disease involvement
Visceral 669 0.55 (0.45–0.67)
Nonvisceral 322 0.96 (0.71–1.30)
Age group
<65 853 0.62 (0.52–0.74)
65–74 113 0.88 (0.53–1.45)
≥75 25 3.51 (1.22–10.13)
Race
White 732 0.63 (0.52–0.77)
Asian 180 0.82 (0.57–1.18)
Other 79 0.59 (0.31–1.11)
Gender
Female 986 0.67 (0.57–0.79)
Male 5 0.00 (0.00–NE)
Baseline ECOG PS
0 611 0.61 (0.49–0.77)
1 370 0.76 (0.59–0.98)
Number of disease sites
<3 605 0.60 (0.48–0.75)
≥3 364 0.73 (0.57–0.94)
Prior anthracycline therapy
Yes 605 0.70 (0.57–0.87)
No 386 0.61 (0.47–0.79)
Baseline liver metastases
Yes 405 0.59 (0.45–0.76)
No 577 0.71 (0.57–0.89)
Unknown 9 1.73 (0.11–27.89)
Baseline bone metastases
Yes 399 0.76 (0.58–0.99)
No 570 0.61 (0.49–0.76)
Unknown 22 1.30 (0.21–8.04)
ER and PR status
ER-positive and/or PR-positive 545 0.72 (0.58–0.91)
ER-negative and PR-negative 426 0.56 (0.44–0.72)
Unknown 20 3.49 (0.75–16.30)
Baseline disease measurability
Yes 786 0.62 (0.52–0.75)
No 205 0.91 (0.59–1.42)
Menopausal status
Premenopausal 451 0.70 (0.54–0.90)
Perimenopausal 38 0.49 (0.20–1.22)
Postmenopausal 400 0.68 (0.53–0.87)
Unknown 79 0.55 (0.31–0.99)
Not applicable 23 0.74 (0.26–2.16)
Prior systemic therapy for MBC
Yes 873 0.69 (0.58–0.82)
No 118 0.51 (0.30–0.85)
Prior trastuzumab treatment for MBC
Yes 836 0.67 (0.56–0.81)
No 155 0.62 (0.40–0.95)
0.2
0.5
1
2
5

Table S1. Dose Delays, Reductions, and Discontinuations.

Dose	Toxicity
T-DM1	
3.6 mg/kg IV q3w	Starting dose
Dose delays	• If significant related toxicities (other than those described below) have not recovered to grade 1 or baseline, dose may be delayed up to 42 days from the last dose (if dosing resumes, it may either be at the same dose level or one dose level lower)
First reduction to 3.0 mg/kg IV q3w	• Platelet count <25,000/mm^3 (after recovering to platelet count ≥75,000/mm^3 or baseline) • AST >3 × ULN (without ALT >3 × ULN) and a subsequent increase of total bilirubin to >2 × ULN within 21 days (after recovering to AST ≤2.5 × ULN and total bilirubin to ≤1.5 × ULN, and after consultation with the medical monitor) • AST/ALT >5 × ULN and/or total bilirubin >1.5 × ULN (after recovering to AST/ALT ≤5 × ULN and/or total bilirubin ≤1.5 × ULN or baseline)
Second reduction to 2.4 mg/kg IV q3w	• Platelet count <25,000/mm^3 (after recovering to platelet count ≥75,000/mm^3 or baseline) with T-DM1 3.0 mg/kg IV q3w • AST >3 × ULN (without ALT >3 × ULN) and a subsequent increase of total bilirubin to >2 × ULN within 21 days (after recovering to AST ≤2.5 × ULN and total bilirubin to ≤1.5 × ULN, and after consultation with the medical monitor) with T-DM1 3.0 mg/kg IV q3w • AST/ALT >5 × ULN and/or total bilirubin >1.5 × ULN (after recovering to AST/ALT ≤5 × ULN and/or total bilirubin ≤1.5 × ULN or baseline)

	with T-DM1 3.0 mg/kg IV q3w
Permanently discontinue T-DM1	• Platelet count <25,000/mm^3 with T-DM1 2.4 mg/kg IV q3w • Grade 3 or 4 hematologic event; platelet counts not recovered to ≥75,000/mm^3 or baseline within 42 days of last dose • ALT >3 × ULN and a subsequent increase of total bilirubin to >2 × ULN within 21 days, regardless of dose level • AST/ALT >5 × ULN and/or total bilirubin >1.5 × ULN with T-DM1 2.4 mg/kg IV q3w • AST/ALT >5 × ULN and/or total bilirubin >1.5 × ULN not recovered to AST/ALT ≤5 × ULN and/or total bilirubin ≤1.5 × ULN or baseline within 42 days of last dose • Grade 3 or 4 peripheral neuropathy not resolved to grade ≤2 within 42 days of last dose • Confirmed CHF (grade ≥3 left ventricular systolic dysfunction per NCI CTCAE v3.0) • LVEF <40% (and confirmed with a repeat assessment within 21 days) or decline in LVEF ≥10% for patients whose LVEF falls to ≤45% (and confirmed with a repeat assessment within 3 weeks without recovery to within 10% of baseline)
Capecitabine	
1000 mg/m^2 PO twice daily (total daily dose of 2000 mg/m^2) on days 1	Starting dose

to 14 of each 21-day treatment cycle	
Dose delays	• To allow grade 2 to 4 adverse events to resolve to grade ≤1
First reduction to 75% of the total daily dose	• Second occurrence of a grade 2 adverse event considered to be significant and/or related that resolves to grade ≤1 • First occurrence of a grade 3 adverse event considered to be significant and/or related that resolves to grade ≤1
Second reduction to 50% of the total daily dose	• Second occurrence of a grade 2 adverse event considered to be significant and/or related that resolves to grade ≤1 (if already being given at 75% of the starting dose) • Third occurrence of a grade 2 adverse event considered to be significant and/or related that resolves to grade ≤1 • First occurrence of a grade 3 adverse event considered to be significant and/or related that resolves to grade ≤1 (if already being given at 75% of the starting dose) • Second occurrence of a grade 3 adverse event considered to be significant and/or related that resolves to grade ≤1 • First occurrence of a grade 4 adverse event considered to be significant and/or related that resolves to grade ≤1 (if thought to be in the patient's best interest)
Permanently discontinue capecitabine	• Any occurrence of a grade 2 adverse event considered to be significant and/or related (if already being given at 50% of the starting dose) • Second occurrence of a grade 3 adverse event considered to be

	significant and/or related (if already being given at 50% of the starting dose) • First occurrence of a grade 4 adverse event considered to be significant and/or related (if thought to be in the patient's best interest or if already being given at 50% of the starting dose) • Second occurrence of a grade 4 adverse event considered to be significant and/or related • Grade 2 to 4 adverse event considered to be possibly related and significant that fails to resolve to grade ≤1 • If lapatinib and capecitabine are both delayed more than 42 consecutive days
Lapatinib	
1250 mg/day PO	Starting dose
Dose delays	• Grade ≥2 toxicity that is considered significant and/or related (that resolves to grade ≤1 or baseline)
Reduction to 1000 mg/day	• Second grade ≥2 toxicity that is considered significant and/or related, that recurs after resolving to grade ≤1 or baseline • LVEF that is grade ≥2 per NCI CTCAE v3.0 or that drops below the institution's lower limit of normal (after ≥14 days if the LVEF recovers to normal and the patient is asymptomatic)
Reduction to 750 mg/day	• Severe hepatic impairment (Child-Pugh class C)
Permanently discontinue	• Grade 2 to 4 adverse event considered to be possibly related and significant that fails to resolve to grade ≤1 or baseline

lapatinib	• If lapatinib and capecitabine are both delayed more than 42 consecutive days

ALT denotes alanine aminotransferase, AST aspartate aminotransferase, CHF, congestive heart failure, CTCAE v3.0 Common Terminology Criteria for Adverse Events version 3.0, IV intravenous, LVEF left ventricular ejection fraction, NCI National Cancer Institute, q3w every 3 weeks, PO orally, T-DM1 trastuzumab emtansine, ULN upper limit of normal.

Table S2. Additional Patient Demographic and Baseline Characteristics.

Characteristic	**Lapatinib Plus Capecitabine (N=496)**	**T-DM1 (N=495)**
Median left ventricular ejection fraction, % (range)*	61 (50–88)	62 (50–87)
Measurable disease, n (%)†	389 (78)	397 (80)
Number of metastatic sites, n (%)†		
<3	307 (62)	298 (60)
≥3	175 (35)	189 (38)
Unknown	14 (3)	8 (2)
Duration of trastuzumab treatment, n (%)		
<1 year	212 (43)	210 (42)
≥1 year	284 (57)	285 (58)
Median time since last trastuzumab treatment, months (range)	1.5 (0–98)	1.5 (0–63)

T-DM1 denotes trastuzumab emtansine.

*****Baseline left ventricular ejection fraction as determined by local assessment; data were available for 472 patients in the lapatinib-plus-capecitabine group and 489 patients in the T-DM1 group.

†Measurable disease and number of metastatic sites at baseline were determined by the independent review committee.

Table S3. Progression-free Survival by Independent Review and by Investigator, and Sensitivity Analysis Results.

	Median PFS, months			
Analysis	**Lapatinib Plus Capecitabine**	**T-DM1**	**HR (95% CI)**	**Log-rank P value**
PFS by independent review				
Stratified analysis	6.4	9.6	0.65 (0.55–0.77)	<0.0001
Unstratified			0.66 (0.56–0.78)	<0.0001
PFS by investigator				
Stratified analysis	5.8	9.4	0.66 (0.56–0.77)	<0.0001
Unstratified			0.66 (0.57–0.78)	<0.0001
Sensitivity analysis censoring for non-protocol therapy				
Stratified analysis	6.7	9.5	0.68 (0.57–0.81)	<0.0001
Unstratified			0.69 (0.58–0.82)	<0.0001

CI denotes confidence interval, HR hazard ratio, PFS progression-free survival, T-DM1 trastuzumab emtansine.

RECIST AND MODIFICATIONS

The following table compares the published RECIST, published by P. Therasse et. al. in JNCI 220, 92:205-16: with the modified RECIST that was utilized for the assessment of response and related parameters throughout the trial. The described modifications represent adaptations of the published criteria based on current radiology and oncology practices, and subsequently provide a more objective and reproducible response assessment. A rationale for the modified criteria is provided as well.

	Original RECIST	Modified Criteria	Modification Rationale
Measurability of Tumor Lesions at Screening	At baseline, tumor lesions will be categorized as follows: measurable (lesions that can be accurately measured in at least one dimension [longest diameter to be recorded] as ≥ 20 mm with conventional techniques or as ≥ 10 mm with spiral CT scan or nonmeasurable (all other lesions, including small lesions [longest diameter < 20 mm with conventional techniques or <10 mm with spiral CT scan] and truly nonmeasurable lesions). Lesions considered to be truly nonmeasurable include the following: bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusion, inflammatory breast disease, lymphangitis cutis/pulmonis, abdominal masses that are not confirmed and followed by imaging techniques, and cystic lesions.	At screening, tumor lesions will be categorized as follows: On spiral CT, for images with lesions with a reconstruction interval of less than or equal to 5 mm, the minimum **measurable** lesion size will be 10 mm; if the reconstruction interval on spiral CT is greater than 5 mm, the minimum lesion size will be double the reconstruction interval. On conventional CT or MRI, for images with lesions with a reconstruction interval of less than or equal to 10 mm, the minimum measurable lesion size will be 20 mm; if the reconstruction interval on conventional CT or MRI is greater than 10 mm, the minimum lesion size will be double the reconstruction interval. **Nonmeasurable** lesions will include all other lesions, including small lesions and truly nonmeasurable lesions. Brain imaging acquired at screening or follow-up or an unscheduled timepoint will undergo radiology review. Brain lesions will be assessed as non-target lesions. Any brain lesions identified by the investigator sites will be taken into consideration by the oncologist in his/her assessment.	Appendix I in the RECIST article - Specifications for Radiologic Imaging / Specific Notes. This allows sites that are capable of performing high quality conventional and spiral CTs or MRIs to participate in the study by allowing double the slice thickness regardless of methodology. Per agreement with Sponsor.
Recording tumor measurements	All measurements should be recorded in metric notation by use of a ruler or calipers. All baseline evaluations should be performed as closely as	All tumor measurements will be recorded in **millimeters** using **electronic** calipers.	To be consistent in the database.

	possible to the beginning of treatment and never more than 4 weeks before the beginning of treatment.		
Selecting target lesions in previously irradiated areas	Tumor lesions that are situated in a previously irradiated area might or might not be considered measurable, and the conditions under which such lesions should be considered must be defined in the protocol when appropriate.	The radiologists **may select** target lesions in **previously irradiated** areas, as radiographically apparent.	RECIST states that a rule must be defined for selecting target lesions in previously irradiated areas.
Specifications by Methods of Measurements	The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline and during follow-up. Imaging based evaluation is preferred to evaluation by clinical examination when both methods have been used to assess the antitumor effect of a treatment.	**None**	**None**
Clinical Examination	Clinically detected lesions will only be considered measurable when they are superficial (e.g., skin nodules and palpable lymph nodes). For the case of skin lesions, documentation by color photography – including a ruler to estimate the size of the lesion – is recommended.	Radiologist may not select target lesions from clinical sources even if no radiographic target lesions are present as determined by the radiologists. The oncologist will incorporate physical exam findings as assessed by the investigators that were not radiographically assessed .They will not select target lesions from clinical sources, and rather will assess qualitatively.	The radiologists will limit measurement of target lesions to CT and MRI scans, the best currently available and most reproducible methods.
Chest X-Rays	Lesions on chest x-ray are acceptable as measurable lesions when they are clearly defined and surrounded by aerated lung. However, CT is preferable.	Chest x-rays **may** undergo review by the radiologists. However, lesions seen on chest x-rays will not be considered measurable and will be followed qualitatively as non-target lesions.	CT and MRI are the best currently available and most reproducible methods for measuring target lesions selected for response assessment.
CT and MRI	CT and MRI are the best currently available and most reproducible methods for measuring target lesions selected for response assessment. Conventional CT and MRI should be performed with contiguous cuts of 10 mm or less in slice thickness. Spiral CT should be performed by use of a 5 mm contiguous reconstruction algorithm; this	CT and MRI will be used as per RECIST. Recommended scanning parameters for this protocol such as slice thickness and reconstruction interval are specified in the Image Acquisition Guidelines.	Appendix I in the RECIST article - Specifications for Radiologic Imaging / Specific Notes. This allows sites that are capable of performing high quality conventional and spiral CTs or MRIs to participate in the

	specification applies to the tumors of the chest, abdomen, and pelvis, while head and neck tumors and those of the extremities usually require specific protocols.		study by allowing double the slice interval regardless of methodology.
Ultrasound	When the primary endpoint of the study is objective response evaluation, ultrasound should not be used to measure tumor lesions that are clinically not easily accessible. It may be used as a possible alternative to clinical measurements for superficial palpable lymph nodes, subcutaneous lesions, and thyroid nodules. Ultrasound might also be useful to confirm the complete disappearance of superficial lesions usually assessed by clinical examination.	Ultrasound **will not be used** to measure tumor lesions.	CT and MRI are the best currently available and most reproducible methods for measuring target lesions selected for response assessment. Ultrasound is necessarily subjective.
Endoscopy/ Laparoscopy	The utilization of these techniques for objective tumor evaluation has not yet been fully or widely validated. Their uses in this specific context require sophisticated equipment and a high level of expertise that may be available only in some centers. Therefore, utilization of such techniques for objective tumor response should be restricted to validation purposes in specialized centers. However, such techniques can be useful in confirming complete histopathologic response when biopsy specimens are obtained.	Endoscopy and laparoscopy **will not be used** to measure tumor lesions.	The utilization of these techniques for objective tumor evaluation has not yet been fully or widely validated.
Tumor Markers	Tumor markers alone cannot be used to assess response. However if markers are initially above the upper normal limit, they must return to normal levels for a patient to be considered in complete clinical response when all tumor lesions have disappeared.	Tumor markers will not be assessed by the independent reviewers.	Per protocol
Cytology and Histology	Cytologic and histologic techniques can be used to differentiate between partial response and complete response in rare cases (e.g., after treatment to differentiate between residual benign lesions and residual malignant lesions in tumor types such as germ cell tumors). Cytologic confirmation of the neoplastic nature of any effusion that appears or worsens during treatment is required when the	In the face of an enlarging effusion/ascites with no progressive non-target disease elsewhere, the radiologist will record tumor response for non-target lesions as Unknown. The overall response for these timepoints will not be driven or altered by this UNK, but will be determined per the rules in the table in Overall Response Section. New pleural effusion/ascites will be recorded as a	Clarification in RECIST regarding new or enlarging pleural effusion or ascites due to ambiguity in original RECIST.

	measurable tumor has met criteria for response or stable disease. Under such circumstances, the cytologic examination of the fluid collected will permit differentiation between response or stable disease (an effusion may be a side effect of the treatment) and progressive disease (if the neoplastic origin of the fluid is confirmed). New techniques to better establish objective tumor response will be integrated into these criteria when they are fully validated to be used in the context of tumor response evaluation.	new lesion. In the case of new effusion/ascites without progressive disease elsewhere, the radiologist will record tumor response of non-target lesions as well as the overall response for the timepoint based on other observable response or progression of disease. In addition, the radiologist will record a comment in the Timepoint Comments section describing the presence, location and any other relevant information about the new effusion/ascites. The oncologist will assign a response according to clinical data (e.g., cytological results). If there are insufficient clinical data available to support a benign condition, then the oncologist will assume malignancy. If the cytology report is missing or unavailable, the oncologist will assess a new or enlarging pleural effusion/ascites as PD.	
Tumor Response Evaluation: *Assessment of overall tumor burden and measurable disease (Baseline)*	To assess objective response, it is necessary to estimate the overall tumor burden at baseline to which subsequent measurements will be compared. Only patients with measurable disease at baseline should be included in protocols where objective tumor response is the primary end point. Measurable disease is defined by the presence of at least one measurable lesion. If the measurable disease is restricted to a solitary lesion, its neoplastic nature should be confirmed by cytology/histology.	Subjects must have either measurable (per RECIST) or non-measurable locally recurrent or metastatic disease. There is no minimum number of target lesions to be identified by the radiologists at screening. If there is no target lesion identified, then the non-target lesions and the appearance of new lesions would be used to evaluate tumor response at post-screening timepoints.	Per protocol all subjects included in the study will be assessed, even if there are no measurable/target lesions as assessed by the radiologists.
Tumor Response Evaluation: Screening documentation of "target" and "nontarget" lesions	All measurable lesions up to a maximum of 5 lesions per organ and 10 lesions in total, representative of all involved organs, should be identified as target lesions and recorded and measured at baseline. Target lesions should be selected on the basis of their size (those with the longest diameter) and their suitability for accurate repeated measurements. A sum of the longest diameter for all target lesions will be calculated and	**Target Lesion Boundary Rules** The primary radiologist reviewers should make every effort to measure (quantitatively assess) all target lesions at post-screening timepoints in spite of imaging of suboptimal quality or poorly defined lesion boundaries. If the lesion has a hypervascular component, that component must be included in the measurement.	RECIST are objective criteria so to the extent possible this minimizes qualitative assessment of target lesions. Hypervascular tissue is viable tumor

	reported as the baseline sum longest diameter. The baseline sum longest diameter will be used as the reference by which to characterize the objective tumor response. All other lesions (or sites of disease) should be identified as nontarget lesions and should also be recorded at baseline. Measurements of these lesions are not required, but the presence or absence of each should be noted throughout the follow-up.	The hypervascular component will continue to be measured in subsequent studies. **Target Lesion Measurement Rules** The primary radiologist reviewers will perform tumor measurements on contrast enhanced CT scans of the chest, abdomen and pelvis at screening. Liver lesions visible by CT will preferably be measured on portal venous phase images. Tumor measurements in the abdomen and pelvis may be performed on MR images if iodine contrast is medically contraindicated. In case of MRI, measurements will be preferably performed in the axial (transverse) plane on contrast enhanced T1 weighted images.	Viable tumor is the portion that is enhancing. This clarifies how to best image enhancement.
Tumor Response Evaluation: *Baseline documentation of "target" and "nontarget" lesions, cont.*		Lesions should be measured using similar images/series throughout the duration of the studies (i.e., lung window CT images, portal venous phase CT images, post-contrast axial T1 MRI images). However, if there is a change from CT to MRI for a given subject at any time during the study, the reviewer will continue to measure provided axial images and that the difference in slice interval is within 5 – 7 mm. Choose the slice where the target lesion is largest at screening. Choose the slice where the longest diameter is largest at follow up, even if it is different from screening. Use all tools available to help measure the lesion (e.g. magnification tools, window/level options in Alice™). The longest diameter of the lesion should be measured even if the actual axis is different from the one used to measure the lesion at screening (or at different timepoints during follow-up). Continue to track and measure target lesions even if the longest diameter of a certain lesion has	For consistency in measurements across visits within subject. Of the many slices to choose from, the slice with the longest in-plane diameter should be chosen. Improves radiologist's accuracy EORTC – RECIST Questions and Answers (www.EORTC.be/recist/) EORTC – RECIST Questions and Answers (www.EORTC.be/recist/)

		fallen below the measurability requirement at screening.	
Tumor Response Evaluation: *Baseline documentation of "target" and "nontarget" lesions, cont.*		If a target lesion becomes less than 5 mm, but is still clearly present, a measurement of 5 mm will be assigned to the longest diameter and the SLD of target lesions will continue to be generated. For any lesion greater than 5 mm, the posted measurement will be retained and used for calculations. If a lesion separates to form discrete lesions on a subsequent study, the longest diameter of each lesion will be calculated and reported separately. • The "child" lesion(s) will be identified with a letter next to the "parent" number, e.g., if lesion #3 splits into two, then the new lesions will be labeled as #3 and #3a. In the event that initially separate lesions have become confluent, the longest diameter of the resulting lesion(s) will be calculated. • The resulting longest diameter will be recorded under one of the original target lesions. Zero mm measurements will be entered for the other target lesion(s) and pertinent comments recorded.	Lesions that are too small may compromise the ability to accurately place electronic calipers for measurements. Five mm is a reasonable estimate of the lower resolution limit of cross sectional imaging techniques. EORTC – RECIST Questions and Answers (www.EORTC.be/recist/) To facilitate tracking EORTC – RECIST Questions and Answers (www.EORTC.be/recist/) To facilitate tracking
Response Criteria: *Target Lesions*	Evaluation of target lesions: This section provides the definitions of the criteria used to determine objective tumor response for target lesions. The criteria have been adapted from the original *WHO Handbook* (WHO handbook for reporting results of cancer treatment. Geneva [Switzerland]: World Health Organization Offset Publication No. 48; 1979), taking into account the measurement of the longest diameter only for all target lesions: complete response – the disappearance of all target lesions; partial response	The radiologist will have the capacity to select the target lesion assessment independently from the application's computations when the measurements do not accurately reflect tumor response. However, the reviewer will be required to enter a comment stating the reason for his/her assessment in these instances. Progressive disease will only be declared when the evidence is unequivocal.	Very small changes in measurements near the limit of imaging resolution, or measurements of normal lymph nodes should not force the radiologist to make inappropriate tumor response assessments. This allows the reviewer to also base the assessment on radiological judgment rather than solely on computational results

	– at least a 30% decrease in the sum of the longest diameter of target lesions, taking as reference the baseline sum longest diameter; progressive disease – at least a 20% increase in the sum of the longest diameter of target lesions, taking as reference the smallest sum longest diameter recorded since the treatment started; stable disease – neither sufficient shrinkage to qualify for partial response nor sufficient increase to qualify for progressive disease, taking as reference the smallest sum longest diameter since the treatment started.	Every effort must be made to measure all target lesions at post-screening timepoints. If a target lesion cannot be measured because of incomplete imaging (i.e., missing anatomical areas, missing slices from examinations, missing one or more films, etc.), poor image quality, or because the lesion has been removed surgically, the target lesion assessment for that timepoint will be limited to "Unknown" or "Progressive disease." If the SLD is indicative of progressive disease, then PD will be specified for target lesion assessment. Otherwise, the SLD will be disregarded and the target lesion assessment will be "Unknown." At following timepoints, when possible, the lesion can again be measured quantitatively, and all overall assessment options are once again valid. See Missing Imaging Data Section for additional information.	in cases where minimal lesion changes may not accurately reflect tumor response. Calculations are incomplete unless all target lesions are measured, except in the case of progression of existing/measured target lesions (increase in SLD in comparison to nadir).
Response Criteria: *Non-Target Lesions*	**Evaluation of nontarget lesions:** This section provides the definitions of the criteria used to determine the objective tumor response for nontarget lesions: complete response – the disappearance of all nontarget lesions and normalization of tumor marker level; incomplete response/stable disease – the persistence of one or more nontarget lesion(s) or/and maintenance of tumor marker level above the normal limits; and progressive disease –unequivocal progression of existing nontarget lesions. ***(Note:*** Although a clear progression of non-target lesions only is exceptional, in such circumstances, the opinion of the treating physician should prevail and the progression status should be confirmed later	Unequivocal progression of non-target lesions will be confirmed by central review, not the treating physician. Non-target lesions will be qualitatively and collectively assessed throughout follow-up. Changes on each non-target lesion group (by anatomical location) will be recorded as: • Complete Response • Incomplete Response/Stable Disease • Progressive Disease • Unknown Unequivocal progression of non-target lesions (i.e., massive growth or enlargement) will be determined qualitatively.	The treating physician may be biased by clinical consideration. Unequivocal progression is not well defined in RECIST.

	by the review panel [or study chair]).	In the case of non-target lesions not imaged, poorly imaged, or because the lesion has been removed surgically, the assessment of non-target lesions will be "Unknown," unless unequivocal progression of evaluable non-target lesions is identified. **Bone Lesions** Any bone imaging that is received from a site will be reviewed. Changes on preexisting bone scan lesions will only have influence over progression if clinical data were acquired. Skeletal survey may be acquired at screening if bone scan is not possible. During the assessment of follow up bone scans, with or without correlative imaging, only the presence of new lesion(s) and site(s) of disease will be noted. **Guidelines on Assessing Bone Scans** **A. Categorization of Bone Scan Lesions** Category I: Lesions on bone scans that are consistent with metastatic disease (with or without supportive imaging studies): • Fusiform/expansile lesion (expansile = beyond boundaries of bone) in the ribs • Uptake involving a large segment of a rib • Hot spot in the pelvis and/or skull not consistent with Paget's disease • Focus of uptake in the scapula (except at acromioclavicular joint) Category II: Lesions on bone scans that are not consistent with metastatic disease (with or without supportive imaging and clinical data): • Focus of uptake in the anterior	"Unknown" will be used for exceptional cases where insufficient data exist, unless progression of evaluable non-target lesions is detected.

		rib/costochondral junction • Focal spot in location consistent with benign condition (specifically in the extremities distal to the mid-humerus and mid-femur) • Hot spot in the pelvis and/or skull consistent with Paget's disease Category III: Lesions on bone scans that are not definitive and may warrant parallel interpretation with other radiographic studies (e.g., x-ray, CT, MRI) or clinical data: • Traumatic fracture, infectious, or inflammatory process. • Focus of uptake in the spine • Foci of uptake consistent with stress fractures • Single hot spot in proximal femur or proximal humerus. • Focus of uptake in the sternum (except sternoclavicular joint and costo-sternal junctions) – CT acquisition preferred • Hot spot in the clavicle (except at sterno and acromioclavicular joints). **B. Guidelines on Recording Bone Scan Lesions at Baseline** 1. If the baseline bone scan lesion(s) is consistent with metastatic disease (Category I), the lesion will be entered into the analysis form and followed. 2. If the baseline bone scan lesion(s) is not consistent with metastatic disease (Category II) the lesion(s) will not be entered into the analysis form. 3. If the baseline bone scan lesion(s) is not definitive (Category III), correlative imaging (x-ray, CT, or MRI) or clinical data is	

		required to identify the nature of the lesion. If the correlative scan shows that the lesion(s) is not malignant, it will not be entered into the analysis form. If the lesion(s) is confirmed to be malignant, it will be entered into the analysis form and the same modality of correlative scan is required at follow-up timepoints. If there are no correlative radiographic studies available, bone scan lesion(s) will be considered malignant and entered into the analysis form. **C. Guidelines for Assessing Follow-Up Bone Lesions** 1. Changes in the character (density, size) on preexisting bone scan lesions should not be used for determination of disease progression or response. 2. New bone lesions that are consistent with Category I will be considered PD. New lesions that are consistent with Category III will be considered PD if confirmed by correlative imaging modalities (plain x-ray, CT, or MRI) or if clinical data is available and indicates the lesion is malignant. 3. If there are no correlative radiographic studies available, new bone scan lesions that are <u>not definitive</u> will be considered malignant.	
Response Criteria: New Lesions	Not distinctly defined in RECIST; clarifications can be found at www.EORTC.be/recist/	**New Lesions:** New lesions will be recorded separately from target/non-target lesions. Any lesion seen for the first time on follow-up with no screening for comparison will be considered a new lesion.	EORTC – RECIST Questions and Answers (www.EORTC.be/recist/) "Appearance of new lesion as indicator of progression is only relevant for overall response evaluation."

<table>
<tr>
<td></td>
<td></td>
<td>New pleural effusion/ascites will be recorded as a new lesion but will not result in an overall response of PD for the timepoint. The radiologist will record a comment in the Timepoint Comments section describing the presence, location and any other relevant information and will override an overall response of PD for the timepoint if it is due solely to a new pleural effusion/ascites. The comments will be available to the oncologist during his/her review.</td>
<td>Guidance from FDA has been to be conservative and assume these lesions are indicative of progressive disease.

EORTC – RECIST Questions and Answers (www.EORTC.be/recist/) “If you are definitely sure on previous images (with the same technique) that this lesion was absent then do not hesitate to conclude progression.”</td>
</tr>
<tr>
<td>Overall Response</td>
<td>Table 1 provides overall responses for all possible combinations of tumor responses in target and nontarget lesions with or without the appearance of new lesions.

<table>
<tr><th>Target Lesions</th><th>Non-Target Lesions</th><th>New Lesions</th><th>Overall Response</th></tr>
<tr><td>CR</td><td>CR</td><td>No</td><td>CR</td></tr>
<tr><td>CR</td><td>IR/SD</td><td>No</td><td>PR</td></tr>
<tr><td>PR</td><td>Non-PD</td><td>No</td><td>PR</td></tr>
<tr><td>SD</td><td>Non-PD</td><td>No</td><td>SD</td></tr>
<tr><td>PD</td><td>Any</td><td>Yes or No</td><td>PD</td></tr>
<tr><td>Any</td><td>PD</td><td>Yes or No</td><td>PD</td></tr>
<tr><td>Any</td><td>Any</td><td>Yes</td><td>PD</td></tr>
</table>

CR = complete response; PR = partial response; SD = stable disease; and PD = progressive disease; IR = incomplete response</td>
<td>If Target Lesion Response is CR, but Non-Target Lesion Response is Unknown and there are no new lesions, the overall timepoint response will be PR. If Target Lesion Response is PR, but Non-Target Lesion Response is Unknown and there are no new lesions, the overall timepoint response will be PR. If Target Lesion Response is SD, but Non-Target Lesion Response is Unknown and there are no new lesions, the overall timepoint response will be SD.

If non-target lesions are UNK due to a missing cytology report, poor quality imaging, or because the lesion has been removed surgically, the overall response will be determined as indicated in the table below.
In the case of missing imaging, see Missing Imaging Data Section.
<table>
<tr><th>Target Lesions</th><th>Non-Target Lesions</th><th>New Lesions</th><th>Overall Response</th></tr>
<tr><td>CR</td><td>CR</td><td>No</td><td>CR</td></tr>
<tr><td>CR</td><td>IR/SD</td><td>No</td><td>PR</td></tr>
<tr><td>CR</td><td>UNK</td><td>No</td><td>PR</td></tr>
</table>
</td>
<td></td>
</tr>
</table>

<table>
<tr><td></td><td></td><td>
<table>
<tr><td>PR</td><td>Non-PD</td><td>No</td><td>PR</td></tr>
<tr><td>PR</td><td>UNK</td><td>No</td><td>PR</td></tr>
<tr><td>SD</td><td>Non-PD</td><td>No</td><td>SD</td></tr>
<tr><td>SD</td><td>UNK</td><td>No</td><td>SD</td></tr>
<tr><td>PD</td><td>Any</td><td>Yes or No</td><td>PD</td></tr>
<tr><td>Any</td><td>PD</td><td>Yes or No</td><td>PD</td></tr>
<tr><td>Any</td><td>Any</td><td>Yes</td><td>PD</td></tr>
<tr><td>UNK</td><td>Non-PD</td><td>No</td><td>UNK</td></tr>
<tr><td>UNK</td><td>PD</td><td>Yes or No</td><td>PD</td></tr>
</table>
</td><td></td></tr>
<tr><td>Overall Response, cont.</td><td>See above</td><td>Overall response for cases with no measurable / target disease at screening as assessed by the radiologists (non-measurable/non-target disease only) will be determined by the following criteria.

<table>
<tr><th>Non-Target Lesions</th><th>New Lesion(s)</th><th>Overall Response</th></tr>
<tr><td>CR</td><td>No</td><td>CR</td></tr>
<tr><td>Incomplete Response/SD</td><td>No</td><td>SD</td></tr>
<tr><td>PD</td><td>Yes or No</td><td>PD</td></tr>
<tr><td>Any</td><td>Yes</td><td>PD</td></tr>
</table>

Overall response for cases with no disease at screening as assessed by the radiologists will be determined by the following criteria.

<table>
<tr><th>New Lesion(s) at Follow-up</th><th>Overall Response</th></tr>
<tr><td>No</td><td>UNK</td></tr>
<tr><td>Yes</td><td>PD</td></tr>
</table>
PD will only be recorded based on unequivocal evidence of progressive disease.</td><td>Proposed criteria based on tumor responses in non-target lesions with or without the appearance of new lesions.

Proposed criteria based on no disease at screening.</td></tr>
<tr><td>Missing Imaging Data</td><td>Not distinctly defined in RECIST</td><td>Missing Imaging Data
Any lesions that are observed at a follow-up imaging timepoint in an area for which there was no corresponding anatomy at baseline or an</td><td></td></tr>
</table>

<table>
<tr>
<td></td>
<td></td>
<td>incomplete imaging timepoint at baseline shall be assumed to represent progressive disease.
In the case of missing or incomplete follow-up imaging, the only possible overall response determination is PD or UNK. This rule regarding missing imaging trumps assessments based on the table in the Overall Response Section.
For assessments having a bone scan only at the current timepoint, if there was soft tissue disease identified at baseline, the radiologist will assign UNK for overall response, unless progressive disease is identified on the bone scan.</td>
<td></td>
</tr>
<tr>
<td>Evaluation of Best Overall Response</td>
<td>The best overall response is the best response recorded from the start of the treatment until disease progression/recurrence (taking as reference for progressive disease the smallest measurements recorded since the treatment started). In general, the subject's best response assignment will depend on the achievement of both measurement and confirmation criteria.</td>
<td>The oncologist will determine best overall confirmed response. Confirmation of response (PR or CR) should be taken into consideration for best overall confirmed response. Confirmation of response will be determined based on the presence of 2 consecutive response determinations, which are at least 4 weeks apart.
<table>
<tr><th>1st Timepoint Response* Assessment</th><th>2nd Timepoint Response Assessment</th><th>Best Overall Confirmed Response</th></tr>
<tr><td>CR</td><td>CR</td><td>CR</td></tr>
<tr><td>CR</td><td>No further evaluation</td><td>SD</td></tr>
<tr><td>CR</td><td>UNK**</td><td>SD</td></tr>
<tr><td>CR</td><td>PD</td><td>PD</td></tr>
<tr><td>PR</td><td>CR</td><td>PR</td></tr>
<tr><td>PR</td><td>PR</td><td>PR</td></tr>
<tr><td>PR</td><td>SD***</td><td>SD</td></tr>
<tr><td>PR</td><td>No further evaluation</td><td>SD</td></tr>
<tr><td>PR</td><td>UNK**</td><td>SD</td></tr>
<tr><td>PR</td><td>PD</td><td>PD</td></tr>
<tr><td>SD</td><td>CR</td><td>SD</td></tr>
<tr><td>SD</td><td>PR</td><td>SD</td></tr>
</table>
</td>
<td>Per agreement with sponsor.</td>
</tr>
</table>

		SD / SD / SD SD / PD / SD SD / UNK / SD SD / No further evaluation / SD UNK / PD / PD UNK / No further evaluation / UNK UNK=Unknown	
Evaluation of Best Overall Response, cont.		* The Best Overall Confirmed Response, other than PD, can only be made after the patient is on-study for a minimum of 6 weeks (42±7 days) from first dose. If a patient only has tumor assessments within this minimal time period (35 days), the patient will have a Best Overall Confirmed Response of Unknown (UNK) unless PD is assessed prior to day 35, in which case the patient's Best Overall Confirmed Response will be PD. ****** Subsequent documentation of CR (or PR) may provide confirmation of previously identified CR (or PR) for patients whose 2nd timepoint response assessment is UNK; if the 3rd timepoint response assessment is CR (or PR) then the Best Overall Confirmed Response will be CR (or PR), for example, PR UNK PR = PR, CR UNK CR = CR. If subsequent timepoint response assessment after a UNK is PD, the Date of Progression will be the date that the PD was first assessed. ***Timepoint Response is SD if there is neither sufficient shrinkage compared to baseline to qualify for partial response nor sufficient increase to qualify for progressive disease, taking as reference the smallest sum longest diameter since	

		the treatment started. At the time of data export, if the BOR is SD or PD, the value will be converted to No Objective Response (NOR).	
Confirmation	The main goal of confirmation of objective response in clinical trials is to avoid overestimating the response rate observed. This aspect of response evaluation is particularly important in nonrandomized trials where response is the primary endpoint. In this setting, to be assigned a status of partial or complete response, changes in tumor measurements must be confirmed by repeat assessments that should be performed no less than 4 weeks after the criteria for response are first met.	<u>Confirmation criteria:</u> In order for a patient to be assigned a status of PR or CR, changes in tumor measurements must be confirmed by repeat assessments that should be performed no less than 4 weeks after the criteria for response are first met. For patients who had bone lesion(s) present at baseline • If the majority of disease was in the soft tissue with little or no bone involvement, x-ray is sufficient to show status of bone lesions at time of response. • Bone scan is only required to confirm CR not PR. For patients with little soft tissue disease and mostly bone involvement or with bone-only disease, a bone scan is needed to confirm response. For patients who had brain lesion(s) present at baseline • A brain scan is necessary to confirm a complete response (CR). • A brain scan is not necessary to confirm a partial response (PR).	Per Protocol.
Reporting of Results	All patients included in the study must be assessed for response to treatment, even if there are major protocol treatment deviations or if they are ineligible. Each patient will be assigned one of the following categories: 1) complete response, 2)	To the extent possible, a response determination should be made. In the case that the reviewer cannot adequately assess tumor response, the timepoint will be labeled as “Unknown” and the reviewer shall provide a concise explanation on	The reviewer should make a determination relying on available information. “Unknown” will be used for exceptional cases where

	partial response, 3) stable disease, 4) progressive disease, or 9) unknown (not assessable, insufficient data). (*Note:* By arbitrary convention, category 9 usually designates the "unknown" status of any type of data in a clinical database.)	the analysis form (except as noted in the Overall Response Section). The only possible overall tumor response determination in the case of incomplete follow-up imaging is PD or Unknown (except as noted in the Overall Response Section).	insufficient data do not support an overall tumor response determination. EORTC – RECIST Questions and Answers (www.EORTC.be/recist/)